VARIABILITY IN FACTOR STRUCTURES OF CLINICIANS' PERSONALITY RATINGS

E.K. ERIC GUNDERSON

EDGAR L. KAPFER

Report Number: 64-23

U. S. NAVY MEDICAL
NEUROPSYCHIATRIC RESEARCH UNIT
SAN DIEGO, CALIFORNIA 92152

BUREAU OF MEDICINE AND SURGERY NAVY DEPARTMENT
WASHINGTON, D.C. 20390

Variability in Factor Structures of Clinicians' Personality Ratings

E. K. Eric Gunderson, Ph. D.

and

Edgar L. Kapfer

United States Navy Medical Neuropsychiatric Research Unit

San Diego, California 92152

Report Number 64-23, October 1964, supported by the Bureau of Medicine and Surgery, Department of the Navy, under Research Task MR005.12-2004, Subtask 1.

Abstract

Personality ratings by military psychologists and psychiatrists who independently examined Navy and civilian volunteers for the U.S. Antarctic Research Program were factor analyzed by professional group and by individual. Factor patterns and trait meanings generally were found to be consistent among psychologists using a personality rating schedule which contained common trait adjectives and somewhat less consistent on another schedule which included specialized Rorschach terminology. Psychologists and psychiatrists as groups were very closely matched on two of three factors; the only notable difference between professional groups in interpretation of the traits utilized was on the Assertive item. Results were considered encouraging for those concerned with construction, standardization, and definition of personality concepts, but further study is needed of the correlates of idiosyncratic frames of reference.

It is a commonplace observation that clinicians differ in their impressions of personality characteristics and dynamics. Clinicians with heterogeneous educational and experiential backgrounds might be expected to vary more widely in ascribed meanings and usage of psychological concepts than those with more homogeneous backgrounds. Military psychologists and psychiatrists, though coming from diverse training institutions and social backgrounds, are exposed to similar professional problems, experiences, and patient populations during their careers. In military clinical practice certain general principles and evaluative frames of reference probably develop as a response to particular features of the military culture. In any case, within a population of military psychologists and psychiatrists it should be possible to delineate the normative or preferred frame of reference for evaluating personality and to determine variability among individual clinicians in relation to the general norm. It would be of theoretical and practical interest to identify idiosyncratic frames of reference, study correlated characteristics of clinicians, and determine consequences for reliability and validity of clinical judgments.

The present study is concerned with the degree of variability that exists among clinicians in their evaluative frames of reference. To what extent will a group of experienced clinicians agree in interpreting and using a set of personality and adjustment ratings? A suitable method for exploring this question would be to have samples of clinical psychologists and psychiatrists evaluate sufficient numbers of persons on a common rating schedule and to factor analyze the ratings of each of the clinicians separately in order to determine the underlying evaluative structure for that individual. The resulting factor patterns could be compared and similarities and differences among them determined. This paper describes an effort to carry out such a study of individual clinicians' evaluative structures.

Methods

Subjects. The subjects for the study were 18 clinical psychologists and the nine psychiatrists who served as examiners for Operation Deep Freeze (U.S. Antarctic Research Program). Psychologists ranged in clinical experience from three to 27 years and psychiatrists from one to 21 years. All examiners had had previous clinical experience with Navy personnel in hospital, dispensaries, or recruit training centers; most of the examiners had had little or no previous knowledge of Operation Deep Freeze or of the volunteer population. About 90 percent of the candidates assessed were Navy personnel who ranged in age from 17 to 38 years and represented a wide variety of military specialties; the remainder of the applicants were civilian technicians or scientists who were evaluated in the same manner as the Navy personnel. Detailed descriptions of the personal and social characteristics of Antarctic volunteers have been presented elsewhere (Gunderson, in press).

Procedures. Each of the candidates was examined and rated independently on a series of personality traits and adjustment predictions by a psychologist and a psychiatrist. Two different rating schedules were employed over the time period covered by the study; the two sets of personality ratings were analyzed separately. Rating variables in both schedules were judged by clinicians to be relevent to adjustment in small isolated Antarctic groups. Candidates were randomly assigned to examiners and generally a single examiner evaluated from 15 to 40 candidates during a single year. Data from applicants rated by the same examiner were accumulated over two or three years for inclusion in the present study.

On the first rating schedule, psychologists rated 24 variables on 4- and 5-point scales. Items were adjectives or sentences descriptive of personality traits and dynamics; the source of clinical impressions was a Rorschach examination. Psychiatrists rated similar items on the first rating schedule, deriving their impressions from a semi-structured clinical interview.

The second rating schedule consisted of 19 variables which were identical for psychologists and psychiatrists. Ratings were made on 6-point scales; again items consisted of common adjectives or brief sentences descriptive of personality traits. For the most part these items differed from those in the first schedule, although seven items in the two rating schedules were essentially the same concepts, a circumstance which helped identify similar factor content. Assessment forms which included the trait ratings generally were filled out immediately after completion of the Rorschach or clinical interview. Assessments were performed either at special screening centers or in Navy hospitals; the same procedures were followed in both settings. Instructions to examiners concerning assessment procedures were general in nature, and no precise definitions of trait scales nor detailed explanation of their use was provided.

Eight psychologists and three psychiatrists were studied in the first series of factor analyses; ten psychologists and six psychiatrists were included in the second series. Ratings were intercorrelated and factor analyzed by the principal axes method for each of the 27 clinicians separately.[2] The analysis extracted a maximum of 10 factors, and rotation of factors was accomplished within the same computer program. The normal varimax method (Kaiser, 1958) was used for an orthogonal rotational solution since this method gives loadings that are less variant over a series of analyses, permitting greater generalization of results. Various objective methods have been proposed for matching factors from one sample to another (Harman, 1960), and a combination of visual inspection and statistical indices were employed in the present study to determine similarities and differences among clinicians' evaluative structures.

Results

A summary of the major results for psychologists on the first rating schedule are shown in Table 1. Results for the three psychiatrists in the first series are omitted from detailed consideration since the psychiatrists' rating schedule differed in important respects from that utilized by psychologists. For simplicity and convenience of interpretation, only the highest loadings for each individual psychologist on his first three factors are indicated. Arabic numerals are used to identify the order or relative magnitude of the individual psychologists' factors; for example, the numeral (1) opposite the Overall Effectiveness item in column A indicates that Psychologist A had a loading of more than .40 for that item on his first (largest) factor; the numeral (2) opposite Assertive-Passive in the same column indicates a loading above .40 for that item on Psychologist A's second factor, and so on. Traits included in the Table are those which loaded most highly on the first three factors (I, II, and III) emerging from a factor analysis of the pooled ratings of all eight psychologists; each factor is represented by three items which form a unified trait cluster or concept. Reading down columns one can discern the unique factor structures of the eight individual psychologists (designated A to H) and the relationship of each evaluative structure to the overall pattern; reading across rows one can see the agreement in clustering of traits and the relative importance (magnitude) of the factors for the various clinicians.

The largest factor emerging from pooled ratings in the first personality rating schedule was best defined by the three most global items in the rating schedule, Overall Effectiveness,

---

[2]The computer program used to obtain the solution was prepared by the staff of the Division of Bio-Statistics, School of Medicine, University of California, Los Angeles, and computer analyses were accomplished by the Statistical Department, U.S. Navy Personnel Research Laboratory, San Diego, California.

Table 1

Variability in Psychologists' Factor Structures: First Personality Rating Schedule

	Psychologist							
Items	A	B	C	D	E	F	G	H
Factor I (All Combined)								
Overall Effectiveness	1[a]	1		3	2	3	2	2
Prefer to Work With	1	1	1	2	2	2	3	2
Likable		1	1	2	2	2		2
Factor II (All Combined)								
Assertive-Passive	2	2	3	3	3	1	1	1
Self-reliant-Dependent	2	2	3	3	3	1	1	1
Modest-Boastful		1	1		3		1	1
Factor III (All Combined)								
Sexual Conflict	3	3	2	1	1	2	2	
Sexual Confusion	3		2					
Castration Anxiety			2	1	1			
N	40	71	88	153	92	54	58	60

[a]Represents a loading of more than .40 on the largest factor for the individual psychologist, 2 represents a loading of more than .40 on the second largest factor, etc.

Likable, and Prefer to Work With. These items together indicate a generally favorable personal impression on the part of the clinician, and the factor has a strongly evaluative connotation. This factor was represented by at least two of the three items in seven of the eight individual analyses; only Psychologist G failed to utilize this cluster. Factor I was not consistently the most important one among individual psychologists, however, in terms of total variance accounted for, since another cluster of items absorbed a larger proportion of the variance for five of the eight clinicians.

The second factor from pooled ratings appeared clearly defined by the items, Assertive-Passive and Self-reliant-Dependent, and to a lesser extent by Modest-Boastful. Agreement was very high among psychologists concerning the salience of this factor since the Assertive and Self-reliant traits form a distinctive pair for all eight psychologists.

The third factor from pooled ratings was concerned with sexual disturbances. It emerged as a distinctive and important factor for four psychologists, as indicated by substantial

loadings on at least two of the three items, and for two of them, D and E, it was the most important factor.

Two other factors, although fairly well-defined in the overall factor analysis, were not consistent over individual factor analyses, and they are omitted from discussion of the results. These factors were composed of the following traits: Absorbs Stress, Emotional Responsiveness, and Adaptable-Rigid; Aggression, Interpersonal Friction, and Hostile-Friendly. Traits which did not have high loadings on any factor in the overall factor analysis nor consistently in the individual factor analyses were: Anxiety, Interpersonal Conflict, Depression, Withdrawn-Sociable, Emotional Control, Tense-Relaxed, Hostile Against Self-Others, Masturbation Concern, and Self-denying-Self-indulgent. These traits were apparently ambiguous, inconsistent, or irrelevant in meaning to these clinicians.

Psychologists' data from the second personality rating schedule is presented in Table 2. The first factor was clearly defined by six rating scales, and almost complete agreement was shown in utilizing these six items as a general evaluative factor. At least five of the six items have high loadings on Factor I' for all ten psychologists.

The second factor for all psychologists combined had high loadings on six traits, and the third factor was defined by three scales. With minor variations in item content, Factors II' and III' were present in the expected order for four of the psychologists (A, B, C, and D) and present but in reversed order for two of the psychologists (E and F). Factor II' did not emerge as a distinct factor for one psychologist (H) and Factor III' failed to emerge for three psychologists (G, I, and J). For all three of the latter individuals, items composing Factor III loaded highly on Factor I'.

Three traits from the second rating schedule, Conforming, Tactful, and Excitable, were not identified with any single factor; one trait, Paranoid, loaded substantially on Factor I' for all psychologists combined but did not consistently load on this factor for individual clinicians.

Factor analyses for the six psychiatrists' personality ratings are shown in Table 3. Essentially the same three factors emerged from factor analyses of the psychiatrists' pooled ratings that had been obtained from the psychologists' ratings. The first factor had the highest loadings on the same six items that defined the psychologists' general evaluative factor. The psychiatrists' second factor contained the same items that defined the psychologists' third factor, namely, Alert, Orderly, and Persevering. In addition, the Assertive

Table 2

Variability in Psychologists' Factor Structures: Second Personality Rating Schedule

	Psychologist									
Items	A'	B'	C'	D'	E'	F'	G'	H'	I'	J'
Factor I' (All Combined)										
Overall Effectiveness	1[a]	1	1	1	1	1	1	1	1	1
Accepted by Peers	1	1	1	1	1	1	1	1	1	1
Accepted by Leaders	1	1	1	1	1	1	1	1	1	1
Prefer to Work With	1	1	1	1	1	1	1	1	2	1
Emotional Control	1	1	1	1		1	1	1	1	1
Flexible	1	1		1	1			1	1	
Factor II' (All Combined)										
Aggressive	2	2	2	2	3	3	2		2	2
Hostile	2	2	2	2	1	3	3	3	2	2
Modest	2			3	3	3	1		3	3
Acts Out	2	2	3	2	3	1	2	2	3	2
Assertive	2	2	2		3	2	2	1	1	2
Hostility Against Self	2	2	2	1	3	3		1	3	2
Factor III' (All Combined)										
Alert	3	3	3		2	2	1	2	1	1
Orderly	3		3	3	2	2	1	2	1	
Persevering		3	3	3	?	2	1	2	1	1
N	41	130	41	68	3?	52	51	34	37	100

[a]Represents a loading of more than .40 on the largest factor for the individual psychologist, 2 represents a loading of more than .40 on the second largest factor, etc.

---

item loaded highly on this factor for psychiatrists. The psychiatrists' third factor included four of the same items that defined the psychologists' second factor and two additional items, Conforming and Tactful. The items Hostile Against Self, Paranoid, and Excitable did not load consistently on a single factor, although Paranoid again loaded substantially on the general evaluative factor for all psychiatrists combined.

Table 3

Variability in Psychiatrists' Factor Structures: Second Personality Rating Schedule

	Psychiatrist					
Items	K	L	M	N	O	P
Factor I" (All Combined)						
Overall Effectiveness	1[a]	2	2	1	1	1
Accepted by Peers	1	1	2	1	1	1
Accepted by Leaders	1	2	1	1	1	1
Prefer to Work With	1	2	2	1	1	1
Emotional Control	1	2	2	1	2	1
Flexibility	1	2	1	1	1	1
Factor II" (All Combined)						
Assertive	2	1	1		1	2
Alert	3	1	1		1	1
Orderly	3	2	1	2	1	3
Persevering	3	1	1		1	1
Factor III" (All Combined)						
Aggressive	2	3	3	3	3	2
Hostile	2	3	3	3	3	1
Modest	2	3	3	2		2
Acts Out	2		3	1		2
Conforming			3	2	2	
Tactful	2	3	3	2	1	1
N	53	70	43	88	55	40

[a]Represents a loading of more than .40 on the largest factor for the individual psychiatrist, 2 represents a loading of more than .40 on the second largest factor, etc.

Discussion

Trait meanings for the general evaluative items and the Assertive-Passive and Self-reliant-Dependent pair of items were generally consistent among psychologists using the first rating schedule which was based solely upon Rorschach examinations. The sex disturbance factor was well-defined in the factor analysis for all raters combined; however, in the individual facto-

analyses items composing this factor tended to have weak loadings or loaded on another factor. This trait cluster was the least well-defined and consistent over individual clinicians of those isolated in the study.

High agreement was shown among the 10 psychologists in interpreting the personality traits utilized in the second rating schedule. Three factors emerged which were represented in the factor patterns of the majority of individual psychologists. Four psychologists employed a somewhat different interpretive scheme from this majority, utilizing only two of the three common factors. Factor I', the general evaluative factor, was very stable and uniform over individual clinicians; Factor II' varied somewhat in item loading patterns among individuals, but three or more items were represented in nine of the 10 individual analyses; Factor III' was least consistent, although at least two of the three items were present in seven of the individual patterns.

Factor patterns for the psychiatrists were somewhat more variable than those for the psychologists on both rating schedules. On the second rating schedule the identical general evaluative factor was present for all psychiatrists that was present for all psychologists. The psychiatrists' second factor was clearly defined for only three individual psychiatrists; for Psychiatrists O and P Factor II" merged with Factor I". For Psychiatrist N this factor does not appear in Table 3 but was not entirely lost since the Assertive and Alert pair of items emerged as the fourth factor for this individual. The third factor was apparent in all six individual factor analyses, although the Aggressive and Hostile item pair split off as a separate factor for two individuals (N and O). The only major differences between overall patterns for psychiatrists and psychologists was the appearance of the Assertive item in different constellations of traits. This trait, which had a definitely negative connotation for psychologists being highly related to Aggressive, Acts Out, Hostile, etc., had a more positive implication in the psychiatrists' lexicon, being part of the Alert, Orderly, and Persevering trait cluster.

A close similarity is noted between the psychologists' Factor II and Factor II' of the present study and Cattell's Dominance-Submissiveness Factor (E) composed of the following traits: Self-Assertive, Confident vs. Submissive, Unsure; Boastful, Conceited vs. Modest-Retiring; Aggressive, Pugnacious vs. Complaisant; Extrapunitive vs. Impunitive, Intropunitive; etc., (Cattell, 1957, p. 108). Also, there is an obvious similarity between Cattell's Super Ego Strength Factor (G), which includes the traits, Persevering, Determined vs. Quitting, Fickle, and Insistently Ordered vs. Relaxed, Indolent (Cattell, 1957, p. 122), and the psychologists'

Factor III' in the second rating schedule. While such comparisons of factor results from data gathered under very different conditions can only be suggestive, they tend to further confidence in the generality of the trait concepts involved.

Agreement between psychologists and psychiatrists as groups with respect to trait meanings obviously was quite good as shown by the similarity in content of item clusters listed in Tables 2 and 3. Agreement also was found to be high using an objective matching procedure described by Harman (1960, p. 257). This index of factor similarity, called by Harman the coefficient of congruence, is analogous to, but is not, a correlation coefficient. The coefficient between first factors (Factor I' and Factor I") for psychologists combined and psychiatrists combined was .90; for psychologists' Factor II' versus psychiatrists' Factor III", the coefficient was .77; for psychologists' Factor III' versus psychiatrists' Factor II", the index was .96. It is inferred that two of the factors were very closely matched between clinician groups and that substantial agreement was present for the third factor. The prominence of the Assertive trait in the psychologists' Factor II' and the absence of this trait in the corresponding psychiatrists' Factor III" could very well account for the lower agreement on that factor.

Differences in factor structures of the two rating schedules can be attributed, in part, to the varying numbers of items included and to the specialized terminology associated with the interpretation of the Rorschach. A deliberate effort was made in constructing the second rating schedule to use only common traits or adjective descriptions since low inter-rater reliabilities had been observed in the earlier schedule, and apparently a greater degree of agreement in meanings of trait concepts was achieved. The number of factors obtained from these rating data was probably limited by the small sampling of traits included in the rating schedules and by the fact that the trait ratings tended to be rather highly inter-correlated. However, the repeated emergence of just three major factors in the present study tends to correspond closely with the number found consistently by Osgood in his work with the "semantic differential." With this technique factor analyses of adjectival scales consistently reduced to three main dimensions, usually labelled Evaluative (large general factor), Potency, and Activity. While most of Osgood's studies provided only average attitudes of many judges, he has summarized the work of one of his students who applied the method to the analysis of individual semantic frameworks (Osgood, 1962). Osgood's technique differed in an important way from the methods of the present study in that his judges rated abstract concepts rather than real persons in an actual clinical setting. Cronbach s (1958) suggestions for mapping

the individual's "implicit personality theory" are believed to offer a particularly fruitful approach for analyses of clinicians' personality concepts.

While there were striking similarities in trait meaning patterns among the majority of clinicians studied, it is noted that Psychologists B, F, G, and H differed from the other psychologists in the first series in omitting sex disturbances as an important factor. Psychologists G', H', I', and J' and Psychiatrists N, O, and P in the second series similarly deviated notably from the modal pattern through merging or splitting of factors. It might be speculated that the absence of shared trait meanings could affect inter-judge reliability; for example, Psychiatrist P's judgments might not be in agreement with nor intelligible to Psychologist F'. In fact, in a series of applicants rated independently by these two clinicians, an average Pearson reliability coefficient of -.03 over all traits in the first factor was obtained. On the other hand, Psychologist F' agreed somewhat better with Psychiatrist M, whose factor pattern closely resembled his own, attaining an average reliability coefficient of .49 over all traits in the first factor. While this example is merely illustrative, it points out a possible consequence of idiosyncratic frames of reference worthy of further investigation, namely, the reduction of inter-judge reliability. Another worthwhile extension of the present research might be to determine the relationships of various characteristics of individual clinicians, such as amount and type of clinical experience, to deviant or unique rating patterns. A practical outcome of the study is the demonstrated feasibility of combining single traits into stable clusters which have similar meanings for the majority of clinicians working in this special setting. Results might be expected to differ in other settings, such as hospitals or out-patient clinics, and in other populations. It seems evident that factor analyses of individual evaluative frameworks may contribute significantly to an understanding of sources of variance in clinical judgments.

References

Cattell, R. B. Personality and Motivation Structure and Measurement. New York: World Book Co., 1957.

Cronbach, L. J. Proposals leading to analytic treatment of social perception scores. In Taquire, R. & Petrullo, L., Person Perception and Interpersonal Behavior. Stanford: Stanford University Press, 1958.

Gunderson, E. K. E. Personal and social characteristics of Antarctic volunteers. Journal of Social Psychology, 1964, 64, 325-332.

Harman, H. H. Modern Factor Analysis. Chicago: University of Chicago Press, 1960.

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 1958, 23, 187-200.

Osgood, C. E. Studies on the generality of affective meaning systems. American Psychologist, 1962, 17, 10-28.

Acknowledgment

Statistical assistance by Mr. F. A. Thompson is gratefully acknowledged. Helpful suggestions were given by Dr. Ardie Lubin.